Characterizing Early-Stage Alzheimer Through Spatiotemporal Dynamics of Handwriting

Christian Kahindo, Mounim A. El-Yacoubi, Sonia Garcia-Salicetti, Anne-Sophie Rigaud, and Victoria Cristancho-Lacroix

***Abstract*—We propose an original approach for characterizing early Alzheimer, based on the analysis of online handwritten cursive loops. Unlike the literature, we model the loop velocity trajectory (full dynamics) in an unsupervised way. Through a temporal clustering based on *K*-medoids, with dynamic time warping as dissimilarity measure, we uncover clusters that give new insights on the problem. For classification, we consider a Bayesian formalism that aggregates the contributions of the clusters, by probabilistically combining the discriminative power of each. On a dataset consisting of two cognitive profiles, early-stage Alzheimer disease and healthy persons, each comprising 27 persons collected at Broca Hospital in Paris, our classification performance significantly outperforms the state-of-the-art, based on global kinematic features.**

***Index Terms*—Alzheimer, clustering of time series, kinematic parameters, online handwriting, probabilistic modeling.**

I. Introduction

ALZHEIMER disease is difficult to detect, at an early stage, owing to its insidious onset. Memory, language, and psychomotor skills are progressively affected. To characterize, on a graphical tablet, the handwriting (HW) of people with early-stage Alzheimer disease (ES-AD), we propose a novel approach by modeling the full HW spatiotemporal dynamics.

Based on online HW, several works have been carried out to characterize Parkinson [1]–[5], Alzheimer [6]–[12], emotion [13], to quantify the effects of drug therapy [14], [15], and for studying the effect of age on HW [16], [17].

A. Conventional HW-Based Alzheimer Assessment

Most works on HW analysis are based on statistical tests [6]–[12]. Few, nonetheless, consider classification models [6], [10], [12]. Both approaches extract from the whole HW task, global parameters as mean velocity [6], [9]–[12] and acceleration [12], mean pressure [6], [12], in-air-time, on-surface-time [6], [10], [12], and task duration [9], [10].

Most statistical test based approaches (ANOVA, etc.) show significant difference between the classes (AD for Alzheimer disease; HC for healthy controls; MCI for mild cognitive impairment) w.r.t. mean velocity [6], [9]. Nonetheless, Yu and Chang [11] did not find any difference between the three groups. Overall, most works claim that patients with cognitive decline have lower velocity, lower pressure, and longer HW time.

In classification approaches, Werner *et al*. [6] obtained, on the three classes, a classification rate from 62.8% to 72%, depending on the task using an LDA with, as input, mean pressure, in-air-time, on-surface-time, in-air-time/on-surface-time ratio, mean velocity, and mini-mental state examination (MMSE) score. Kawa *et al*. [10] applied LDA to classify MCI vs. HC, based on features selected according to the Mann–Whitney *U*-test, and obtained a classification rate of 70%. Garre-Olmo *et al*. [12] also used LDA taking as input a combination of features (pressure, acceleration, speed, complexity, and age) and reached a classification rate from 63.5% to 100%.

Both categories of approaches consider global parameters extracted from the whole HW task (e.g., mean velocity, etc.), which is a disadvantage since the parameters' dynamics are lost. The statistical tests are sometimes inconclusive or even contradictory (e.g., [6] vs. [11] w.r.t. velocity). This may be explained by the relatively poor discriminative power of the global parameters, the nonhomogeneity within the same cognitive profile (e.g., patients with ES-AD do not necessarily show the same HW degradation), or overfitting due to small data sets. Unlike most statistical methods, the classification approaches [6], [10], [12] take as input a set of parameters rather than one, each in turn. The very small training datasets, however, induce a strong overfitting. The reported classification results are often misleading (e.g., [6], [12]) as they are obtained on the very data the models are trained on.

B. Our Approach

We propose a novel technique to characterize ES-AD w.r.t. HC by analyzing the kinematics of online HW, on a task consisting of four cursive $\ell\ell\ell\ell$ series, written by each participant (see Fig. 1). Instead of comparing ES-AD and HC based on global parameters, our approach addresses the limits of the studies mentioned above by modeling the *full* dynamics of these parameters. To do so, we first automatically segment the $\ell\ell\ell\ell$ series into individual loops. To characterize the variability of loops over the two classes, we define a dictionary of prototype

Manuscript received September 25, 2017; revised December 16, 2017; accepted December 17, 2017. Date of publication January 17, 2018; date of current version June 22, 2018. This work was supported in part by the Institut Mines Telecom, in part by the MAIF Foundation, and in part by the Hospital Broca in Paris. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Mehdi Moradi. *(Corresponding author: Mounim A. El-Yacoubi.)*

C. Kahindo, M. A. El-Yacoubi, and S. Garcia-Salicetti are with SAMOVAR, Telecom SudParis, CNRS UMR 5157, Paris Saclay University, Evry 91000, France (e-mail: Christian.Kahindo@telecom-sudparis.eu; mounim.el_yacoubi@telecom-sudparis.eu; Sonia.Garcia@telecom-sudparis.eu).

A.-S. Rigaud and V. Cristancho-Lacroix are with AP-HP, Groupe Hospitalier Cochin Paris Centre, Pôle Gérontologie, Hôpital Broca, Paris 75013, France, and also with EA 4468, Paris Descartes University, Paris 75006, France (e-mail: Anne-sophie.Rigaud@aphp.fr; Victoria.Cristancho@aphp.fr).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2018.2794500

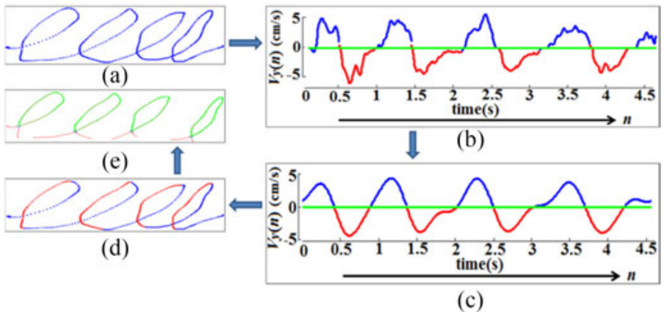

Fig. 1. Loop segmentation: (a) input loops series, (b) the $V_y(n)$ signal, (c) low-pass filtering by the fundamental frequency (here 1 Hz), (d) segmentation into ascending and descending strokes, and (e) extraction of the loops.

(medoid) loops by a temporal clustering scheme over the training data. The segmentation allows to increase the size of the training data, and accordingly the reliability of the clustering; also, individual loop-based clusters are more likely to be homogeneous than would be the clusters of entire $\ell\ell\ell\ell$ series. Our clustering is based on the *K*-medoids algorithm, with a dynamic time warping (DTW) dissimilarity measure that accommodates the sequential aspect of the loops. Each cluster, consists of a set of loops pertaining to the two classes in different proportions, reflecting the cluster power in discriminating the two classes. To leverage all the loops generated by a given person in the test phase, we consider a Bayesian formalism that aggregates the contribution of each loop to make the classification decision over the two classes (ES-AD and HC). Our experimental results, significantly outperform the state-of-the-art. The organization of the letter is the following: Section II presents the major phases of our approach, namely, the segmentation of the series into loops, feature extraction, the temporal clustering of the loops, and the classification of subjects. Section III describes our experiments and results. Finally, in Section IV, we conclude and give some perspectives of our work.

II. Proposed Method

A. Participants

Data collection was carried out at Broca Hospital in Paris. ES-AD patients were diagnosed based on DSM-5 criteria [18], and included if their MMSE [18] was over 20. HC subjects performed neuropsychological tests to ensure their cognitive profile is normal. Participants with medical problems such as stroke and other neurodegenerative diseases were not included. The dataset consists of 54 participants (27 ES-AD and 27 HC), with a mean age of 79.7 ± 6.4 and 73.2 ± 5.7, respectively. All participants freely signed a consent form after receiving information on the study.

B. Equipment

HW was acquired on a WACOM Intuos Pro Large tablet with an inking pen. A paper was fixed on the tablet to allow a natural visual feedback. The tablet records, at a sampling rate of 125 Hz, pen position $\{x(t), y(t)\}$ and pressure $p(t)$ over time, as well as the pen's in-air trajectory up to 2 cm off the tablet surface.

C. Segmentation Into Loops and Feature Extraction

We segment each $\ell\ell\ell\ell$ series into isolated ℓ letters, from which we keep only the loop part for subsequent feature extraction. As each series is a periodical signal, with a period covering an entire l letter, we harness the fast Fourier transform (FFT) for segmentation as follows: a low-pass filter is applied to smooth $V_y(n)$, the vertical velocity signal of the $\ell\ell\ell\ell$ series, by setting the FFT cutoff frequency to the series' fundamental frequency F_0. F_0 is, thus, not only dependent on the writer, but it is on his/her particular $\ell\ell\ell\ell$ series, and whether it is spread out (F_0 is low) or narrow (F_0 is high). We then apply the inverse Fourier transform and segment the trajectory at points *n* where $V_y(n) = 0$. Each loop is then retrieved by merging its two consecutive ascending and descending strokes, a stroke being the segment between two consecutive points with $V_y(n) = 0$ (see Fig. 1). Setting the cutoff to the fundamental frequency allows to segment the loops irrespective of the irregularities and tremors in shaky HW, more reliably than manually based thresholding techniques. The frequency varies in general between 1 and 4 Hz.

We extract, at each loop point, the velocity in the *x*- and *y*-direction. Velocity is the first derivative of position: at each point *n*, the horizontal and vertical velocities $V_x(n)$ and $V_y(n)$ are computed as $V_x(n) = \Delta x(n)/\Delta t(n)$; $V_y(n) = \Delta y(n)/\Delta t(n)$, where $\Delta x(n) = x(n+1) - x(n-1)$, and $\Delta t(n) = t(n+1) - t(n-1)$ [19].

D. DTW-Based K-Medoids Clustering Approach

Each loop is represented by its pointwise velocity $V_x(n)$ and $V_y(n)$ describing the pen trajectory. The dissimilarity between two loops is computed by DTW [20], which matches sequences of different lengths. Clustering is performed by the *K*-medoids algorithm [21], which assigns each sample to its closest medoid. A medoid is the closest loop to all the loops within the same cluster.

E. Bayesian-Based Classification

To model the loops' full dynamics, we discard support vector machine-like classifiers (LDA, etc.) that take as input global parameters like average speed, and pressure, as in the state-of-the-art. We rely instead on the temporal clustering of loops through *K*-medoids-DTW. Each resulting cluster will correspond to some writing style and will consist of loops pertaining either to HC or to ES-AD in different proportions.

To merge the contributions of the different clusters, we consider a Bayesian formalism that takes into account the discriminative power of each. We use Bayes' rule to compute for each person the posterior probability to be ES-AD or HC given his/her respective data (loops). Each person usually produces 16 loops, which are distributed over the clusters obtained by *K*-medoids on the loops' training set. For subject s_i, the posterior probability of class C_k (ES-AD or HC), given his/her data D_i (loops) is computed as follows:

$$P(C_k|D_i) = P(D_i|C_k) \times P(C_k) / P(D_i) \tag{1}$$

where $P(D_i) = \sum_{k=\mathrm{HC},\ \mathrm{ES-AD}} P(D_i/C_k) \times P(C_k)$ and $P(C_k)$ is the *a priori* probability of class C_k (50% in our case). Assuming the N_i loops of person s_i are class-conditionally

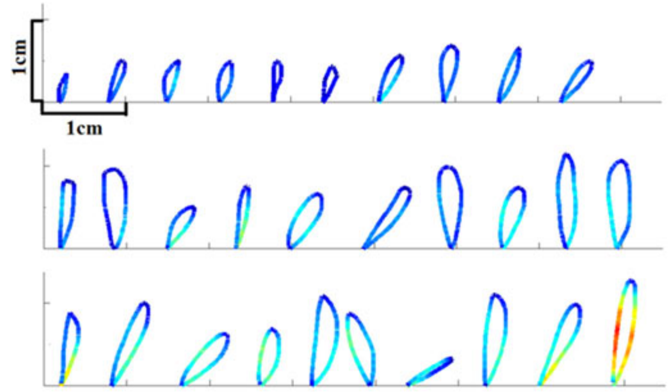

Fig. 2. Medoids characterized by their velocity trajectory, ordered by DTW, with the smallest medoid as reference. The color scale characterizes velocity magnitude: blue stands for low speed and red for high speed.

independent, we have

$$P\left(D_i|C_k\right) = \prod_{j=1}^{N_i} P(B_j^i|C_k) \tag{2}$$

where B_j^i is the closest cluster to the *j*th loop of person s_i. Thus

$$P(B_j^i/C_k) = P(C_k/B_j^i) \times P\left(B_j^i\right)/P\left(C_k\right). \tag{3}$$

$P(B_j^i)$, the *a priori* probability of cluster B_j^i, can be estimated by

$$P\left(B_j^i\right) = N_{B_j^i} / \, N_{\text{Total}} \tag{4}$$

where $N_{B_j^i}$ is the number of loops in cluster B_j^i and N_{Total} is the total number of loops. Likewise

$$P\left(C_k/B_j^i\right) = N_k^j / N_{B_j^i} \tag{5}$$

where N_k^j is the number of loops in cluster B_j^i from class *k* (ES-AD or HC). Each subject s_i producing data D_i is then classified by selecting the class (HC or ES-AD) with the maximum *a posteriori* probability:

$$C_k^* = \underset{k=\text{HC},\ \text{ES-AD}}{\arg\max}\ P(C_k/D_i). \tag{6}$$

III. Experimental Results

We perform *K*-medoids clustering, based on the sequential velocity representation $V_x(n)$,$V_y(n)$, of all the loops from all subjects. The total number of loops is 866, a little more than 864 ($27 \times 2 \times 4 \times 4$), as two persons, produced each, one extra loop, in one of his/her four trials. To get meaningful clusters w.r.t. the size of data, we have tried several numbers of clusters, by varying *K* between 10 and 50, and obtained similar optimal performance for *K* between 30 and 50. Here, we report the results for $K = 30$. The obtained medoids are displayed in Fig. 2. Note the diversity of the 30 medoids, both in terms of shape and dynamics. They are ordered based on their DTW distance to the smallest medoid, taken as reference. This highlights the quality of DTW in matching loops with different velocity dynamics ($V_x(n)$, $V_y(n)$).

Fig. 3 displays examples of clusters, two containing mostly HC subjects [see Fig. 3(a) and (b)] and two mostly ES-AD

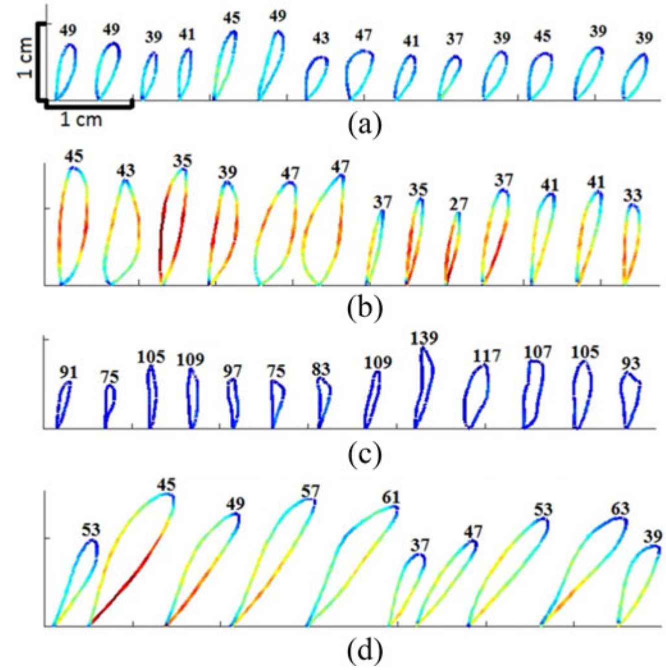

Fig. 3. Samples from four clusters. For each, we report its number of loops, the number of loops for each class, and the number of points for each loop.

patients [see Fig. 3(c) and (d)]. Thanks to our *K*-medoids-DTW clustering, each cluster detects a unique combination of several HW features like the full velocity dynamics, fluidity, shakiness, slant, and size in terms of number of points, which is correlated, but not always, to spatial size.

The loops gathered in the cluster in Fig. 3(a) were written at a moderate speed, with a small-to-average size and a regular shape, while those grouped in the cluster in Fig. 3(b) were written at a high speed, with a large size and a regular shape. These two clusters reflect two different *profiles* of HC writers. For the same reason, the two clusters displayed in the clusters in Fig. 3(c) and (d) represent, by contrast, two different *profiles* of ES-AD patients, the first with loops written at low speeds and showing irregular shapes, and the second with loops of moderate to high speeds and regular shapes. These findings reveal the *nonhomogeneity* of both HC and ES-AD cognitive profiles, which is not taken into account in the current state-of-the-art.

For performance assessment, we consider the *classification* rate, the percentage of correct answers using the Leave-one-person-out procedure. Given our dataset of 54 subjects (27 HC, 27 ES-AD), we consider 54 training runs, in which each subject is taken out in turn for validation, and the remaining 53 for training. Therefore, the validation classification rate is the average of the 54 binary decisions on each held out subject, while the training classification rate is the average of the 54 classification rates obtained, each, on the 53 training subjects remaining for each held-out person. The classification rates are shown in Table I (third column). By exploiting the full dynamics ($V_x(n)$, $V_y(n)$), we obtain a classification rate of 74% on the validation set, which represents an increase of 24% over a random guess (ES-AD and HC represent, each, 50% of the total population), which indicates a relative improvement of about 50%. By comparison, adopting the state-of-the-art scheme by feeding the

TABLE I
CLASSIFICATION RATES IN % ON THE VALIDATION SET

	LDA $(\bar{V}_x, \bar{V}_y)$	Bayes $(\bar{V}_x, \bar{V}_y)$	Bayes $(V_x(n), V_y(n))$
Classification	51.9	64.0 ± 1.0	74.0 ± 3.0

Left: LDA with $(\bar{\boldsymbol{V}}_{\boldsymbol{x}}, \bar{\boldsymbol{V}}_{\boldsymbol{y}})$, and middle and right: Bayes classification with clustering of $(\bar{\boldsymbol{V}}_{\boldsymbol{x}}, \bar{\boldsymbol{V}}_{\boldsymbol{y}})$ and velocity trajectory, respectively.

average velocity $(\overline{V_x}, \overline{V_y})$ as input to LDA, we obtain a classification rate of 51.9%, a bare improvement over a random guess. However, as the two approaches are different in three aspects (Bayesian scheme instead of LDA, clustering or not prior to classification, and full velocity dynamics modeling instead of average velocity), we run an additional experiment to disentangle the three factors mentioned above. To do this, we have performed a clustering of the loops based on their $(\overline{V_x}, \overline{V_y})$ values. The clustering is based on *K*-medoids as before, but now it takes as input $(\overline{V_x}, \overline{V_y})$, with a Euclidian distance as a dissimilarity measure, instead of DTW. Then we apply the Bayesian classification as before. The results, shown in Table I (second column), show a classification rate of 64% on $(\overline{V_x}, \overline{V_y})$, around 12% better than the state-of-the-art based on LDA, but 10% less than our Bayesian scheme based on the velocity full trajectories.

This can be explained in the following way:

1) The clustering characterizes a writer by the distribution of average or trajectory velocity over the different loops s/he produces. A sharp distribution reflects a writer's tendency to write while maintaining the same (average or trajectory) speed rhythm throughout the loops; a flat one reflects, by contrast, his/her tendency to write loops with different (average or trajectory) speeds, as the rhythm is not maintained. This across-loops variability is an important writing style information that is overlooked by all state-of-the-art methods, and we see that considering it in our approach, based on average velocity only, allows a significant increase of 12% (64.0% vs. 51.9%).
2) The actual improvement of modeling the loop's trajectory over the average value is 10% (74.0% vs. 64.0%). This shows that modeling the full velocity dynamics, in an unsupervised way, is significantly better for uncovering HW impairments.

Therefore, the large improvement obtained by our approach results from the combination of clustering and unsupervised trajectory modeling. It is not surprising, then, that average velocity in [11] showed no significant difference between AD and HC, on a similar task (3×8 loops).

For comparison, we assess the main classification techniques used in the literature, LDA. First, we assessed several state-of-the-art average spatiotemporal parameters, taken each separately: average velocity $(\overline{V_x}, \overline{V_y})$, average acceleration $(\overline{A_x}, \overline{A_y})$, average jerk $(\overline{J_x}, \overline{J_y})$, average pressure $\bar{P}$, and total task time *T*. The results on the learning and validation sets are reported in Table II. For benchmarking, we assessed [6] and [12] that considered, each, different tasks. For the task of copying a Hebrew text of 107 characters, Werner *et al.* [6] selected average pressure and on-paper time as input to LDA, and obtained a 59.6% rate of classifying the three classes HC (41 subjects), MCI (31), and AD (22). On a dictated Spanish sentence task, and based on average speed, acceleration, and

TABLE II
LDA-BASED CLASSIFICATION RATES IN % WITH GLOBAL KINEMATIC FEATURES

	$\bar{V}_x$ $\bar{V}_y$	$\bar{A}_x$ $\bar{A}_y$	$\bar{J}_x$ $\bar{J}_y$	$\bar{P}$	T	$\bar{P}$ T [6]	$\overline{\lvert V\rvert}$ $\bar{P}$ $\overline{\lvert A\rvert}$ [12]	$(\overline{V_x}, \overline{V_y})$ $(\overline{A_x}, \overline{A_y})$ $(\overline{J_x}, \overline{J_y})$ $\bar{P}, T$
Training set	55.9	49.8	50.0	50.6	49.7	51.6	57.7	50.4
Test set	51.9	44.0	48.0	35.0	40.7	27.8	46.0	46.3

pressure, Garre-Olmo *et al.* [12] obtained a 100% rate for classifying 17 HC vs. 12 MCI, and 75% for classifying 17 HC vs. $\{12\,\text{MCI} + 23\,\text{AD}\}$. It is important to note that the classification rates and the feature selection in [6] and [12] were obtained on the training set, as no cross validation was considered. In our experiments, we have considered the exact settings in [6] and [12] for our loops series task, but using the Leave-one-person out procedure. These results are reported as well in Table II. Note that we do not report confidence intervals since LDA does not require a random parameter initialization.

Table II confirms that all the average spatiotemporal parameters are, each, basically noise, even on the training set, except for average velocity, which is slightly above 50% (55.9% on training and 51.9% on the validation set). It is not surprising, then, that any combination of them is equivalent to a random choice, with the additional burden that the increase of dimensionality worsens accordingly the overfitting issue. As explained earlier, our huge improvement over these approaches, obtained with velocity only, is due to two key modeling aspects, the clustering that models the across-loops variability for each subject, and the modeling of the full velocity trajectory ($V_x(n)$, $V_y(n)$), which characterizes in a much richer way the subject's writing style.

IV. CONCLUSION AND PERSPECTIVES

We have proposed an original approach for characterizing early Alzheimer, based on the analysis of online handwritten cursive loops. We propose an unsupervised learning approach with the aim of uncovering homogeneous *subgroups* from all the population $\{\text{HC} + \text{ES-AD}\}$, some of which possibly associated with a given cognitive profile. To model the full loops' dynamics, we consider a *K*-medoids clustering of the velocity trajectories of the handwritten loops', with a DTW, as a dissimilarity measure, to accommodate data sequences of different lengths. We use, for classification, a Bayesian framework to combine probabilistically the contributions from each cluster. Based on velocity only, we obtain, on the validation set, a classification rate of 74%, which outperforms the main state-of-the-art classification approaches by more than 50%, even if the latter takes as input not only velocity but other kinematic parameters as well. This shows the power of our approach, which analyzes the full velocity dynamics, instead of global average parameters.

We can expect that modeling the full dynamics of other spatiotemporal parameters like acceleration, jerk, pen pressure, or position, will show good performance as well by uncovering clusters that highlight the writing styles with other view angles. Future work will be focused on assessing these features and their fusion for better classification.

REFERENCES

[1] H. L. Teulings and G. E. Stelmach, "Control of stroke size, peak acceleration, and stroke duration in Parkinsonian handwriting," *Hum. Movement Sci.*, vol. 10, no. 2, pp. 315–334, May 1991.

[2] H. L. Teulings, J. L. Contreras-Vidal, G. E. Stelmach, and C. H. Adler, "Parkinsonism reduces coordination of fingers, wrist, and arm in fine motor control," *Exp. Neurol.*, vol. 146, no. 1, pp. 159–170, Jul. 1997.

[3] V. Gemmert, W. Arend, H. L. Teulings, and G. E. Stelmach, "The influence of mental and motor load on handwriting movements in Parkinsonian patients," *Acta Psychol.*, vol. 100, no. 1, pp. 161–175, Nov. 1998.

[4] P. Drotár, J. Mekyska, I. Rektorová, L. Masarová, Z. Smekal, and M. Faundez-Zanuy, "A new modality for quantitative evaluation of Parkinson's disease: In-air movement," in *Proc. IEEE 13th Int. Conf. Bioinformat. Bioeng.*, 2013, pp. 1–4.

[5] S. Rosenblum, M. Samuel, S. Zlotnik, I. Erikh, and I. Schlesinger, "Handwriting as an objective tool for Parkinson's disease diagnosis," *J. Neurol.*, vol. 260, no. 9, pp. 2357–2361, 2013.

[6] P. Werner, S. Rosenblum, G. Bar-On, J. Heinik, and A. Korczyn, "Handwriting process variables discriminating mild Alzheimer's disease and mild cognitive impairment," *J. Gerontol., Psychol. Sci.*, vol. 61, no. 4, pp. 228–236, 2006.

[7] J. H. Yan, S. Rountree, P. Massman, R. Smith Doody, and H. Li, "Alzheimer's disease and mild cognitive impairment deteriorate fine movement control," *J. Psychiatric Res.*, vol. 42, no. 14, pp. 1203–1212, Oct. 2008.

[8] A. Schröter, R. Mergl, K. Bürger, H. Hampel, H. J. Möller, and U. Hegerl, "Kinematic analysis of handwriting movements in patients with Alzheimer's disease, mild cognitive impairment, depression and healthy subjects," *Dementia Geriatric Cogn. Disorders*, vol. 15, no. 3, pp. 132–142, 2003.

[9] M. J. Slavin, J. G. Phillips, J. L. Bradshaw, K. A. Hall, and I. Presnell, "Consistency of handwriting movements in dementia of the Alzheimer's type: A comparison with Huntington's and Parkinson's diseases," *J. Int. Neuropsychol. Soc.*, vol. 5, no. 1, pp. 20–25, 1999.

[10] J. Kawa, A. Bednorz, P. Stepien, J. Derejczyk, and M. Bugdol, "Spatial and dynamical handwriting analysis in mild cognitive impairment," *Comput. Biol. Med.*, vol. 82, pp. 21–28, Mar. 2017.

[11] N. Y. Yu and S. H. Chang, "Kinematic analyses of graphomotor functions in individuals with Alzheimer's disease and amnestic mild cognitive impairment," *J. Med. Biol. Eng.*, vol. 36, no. 3, pp. 334–343, Jun. 2016.

[12] J. Garre-Olmo, M. Faúndez-Zanuy, K. López-de-Ipiña, L. Calvó-Perxas, and O. Turró-Garriga, "Kinematic and pressure features of handwriting and drawing: Preliminary results between patients with mild cognitive impairment, Alzheimer disease and healthy controls," *Curr. Alzheimer Res.*, vol. 14, no. 9, pp. 960–968, Mar. 2017.

[13] L. Likforman-Sulem, A. Esposito, M. Faundez-Zanuy, S. Clémençon, and G. Cordasco, "EMOTHAW: A novel database for emotional state recognition from handwriting and drawing," *IEEE Trans. Hum.-Mach. Syst.*, vol. 47, no. 2, pp. 273–284, Apr. 2017.

[14] P. C. Poluha, H. L. Teulings, and R. H. Brookshire, "Handwriting and speech changes across the levodopa cycle in Parkinson's disease," *Acta Psychol.*, vol. 100, no. 1, pp. 71–84, Nov. 1998.

[15] M. P. Caligiuri, H. L. Teulings, J. V. Filoteo, D. Song, and J. B. Lohr, "Quantitative measurement of handwriting in the assessment of drug-induced parkinsonism," *Hum. Movement Sci.*, vol. 25, no. 4, pp. 510–522, Oct. 2006.

[16] S. Rosenblum, B. Engel-Yeger, and Y. Fogel, "Age-related changes in executive control and their relationships with activity performance in handwriting," *Hum. Movement Sci.*, vol. 32, no. 2, pp. 363–376, Apr. 2013.

[17] G. Marzinotto *et al.*, "Age-related evolution patterns in online handwriting," *Comput. Math. Methods Med.*, vol. 2016, 2016, Art. no. 3246595.

[18] *Diagnostic and Statistical Manual of Mental Disorders, Fifth ed. (DSM-5)*. Arlington, VA, USA: Amer. Psychiatric Assoc., 2013.

[19] I. Guyon, P. Albrecht, Y. Le Cun, J. Denker, and W. Hubbard, "Design of a neural network character recognizer for a touch terminal," *Pattern Recog.*, vol. 24, no. 2, pp. 105–119, 1991.

[20] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 26, no. 1, pp. 43–49, Feb. 1978.

[21] L. Kaufman and P. J. Rousseeuw, "Clustering by means of Medoids," in *Statistical Data Analysis Based on the L1 Norm and Related Methods*, Y. Dodge, ed. Amsterdam, The Netherlands: North Holland, 1987, pp. 405–416.